DOCUMENT RESUME

ED 211 605 TM 820 061

AUTHOR Tindal, Gerald; Deno, Stanley L.
TITLE Daily Measurement of Reading: Effects of Varying the Size of the Item Pool.
INSTITUTION Minnesota Univ., Minneapolis. Inst. for Research on Learning Disabilities.
SPONS AGENCY Office of Special Education (ED), Washington, D.C.
REPORT NO IRLD-RR-55
PUB DATE Jul 81
CONTRACT 300-80-0622
NOTE 27p.
AVAILABLE FROM Editor, IRLD, 350 Elliot Hall, 75 East River Road, University of Minnesota, Minneapolis, MN 55455 ($3.00).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Elementary Education; Evaluation Methods; *Informal Reading Inventories; *Item Banks; *Reading Achievement; Reading Instruction; Reading Tests; Special Education; *Test Construction; *Test Items
IDENTIFIERS *Test Curriculum Overlap

ABSTRACT

Evidence exists that reading achievement can be measured simply and validly by having students read aloud for one minute from vocabulary lists drawn from their basal reading series. Direct and frequent measurement of student performance using this procedure provides a means for continuously evaluating a student's instructional program. The present study investigated the effects of varying the size of the population of words from which test items for daily testing were sampled. Results indicated that grade-level lists were more sensitive to changes in performance and that across-grade lists produced less variability in performance. The size of the word population did not seem to influence the ability of judges to perform visual analyses of instructional effects. The implications of the findings for measurement and teaching are discussed. (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED211605

University of Minnesota

Research Report No. 55

DAILY MEASUREMENT OF READING: EFFECTS OF VARYING THE SIZE OF THE ITEM POOL

Gerald Tindal and Stanley L. Deno

IRLD

Institute for Research on Learning Disabilities

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

TM 820 061

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

J. Vasseldyke

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

IRLD

Director: James E. Ysseldyke

Associate Director: Phyllis K. Mirkin

The Institute for Research on Learning Disabilities is supported by a contract (300-80-0622) with the Office of Special Education, Department of Education, through Title VI-G of Public Law 91-230. Institute investigators are conducting research on the assessment/decision-making/intervention process as it relates to learning disabled students.

During 1980-1983, Institute research focuses on four major areas:

- Referral
- Identification/Classification
- Intervention Planning and Progress Evaluation
- Outcome Evaluation

Additional information on the Institute's research objectives and activities may be obtained by writing to the Editor at the Institute (see Publications list for address).

The research reported herein was conducted under government sponsorship. Contractors are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent the official position of the Office of Special Education.

Research Report No. 55

DAILY MEASUREMENT OF READING:
EFFECTS OF VARYING THE SIZE OF THE ITEM POOL

Gerald Tindal and Stanley L. Deno

Institute for Research on Learning Disabilities

University of Minnesota

July, 1981

Abstract

Evidence exists that reading achievement can be measured simply and validly by having students read aloud for one minute from vocabulary lists drawn from their basal reading series. Direct and frequent measurement of student performance using this procedure provides a means for continuously evaluating a student's instructional program. The present study investigated the effects of varying the size of the population of words from which test items for daily testing were sampled. Results indicated that grade-level lists were more sensitive to changes in performance and that across-grade lists produced less variability in performance. The size of the word population did not seem to influence the ability of judges to perform visual analyses of instructional effects. The implications of the findings for measurement and teaching are discussed.

Daily Measurement of Reading:

Effects of Varying the Size of the Item Pool

Two activities inherent in instruction are observation of student performance and adjustment of instructional tactics based on those observations. Typically, of course, teachers' observations are informal and tactical adjustments are unsystematically introduced. As the requirements for accountability increase, however, and as instructional designers attempt to improve instructional systems through educational technology, greater emphasis is placed on tests as a basis for observing student performance and evaluating program effectiveness. One effect of increasingly using tests in this way is the misuse of commercially prepared standardized tests (Salvia & Ysseldyke, 1981). Tests designed for psychometric purposes are used as edumetric instruments (Carver, 1974) and poor fits between "what is taught" and "what is tested" occur (Jenkins & Pany, 1978; Skager, 1971).

One alternative to commercially prepared achievement tests is direct observation and recording of student performance within the curriculum (Lovitt, Schaff, & Sayre, 1970; White & Haring, 1980). Assessment of performance within the curriculum in which the student is receiving instruction is an attractive alternative since it reduces the gap between what is taught and what is tested. Further, the use of informal classroom measures may make it possible to tailor measurement to the individual student and the educational program, to measure student performance on a frequent basis, and to monitor and evaluate the effectiveness of instructional programs.

Although the use of informal measures appears helpful in monitor-

ing students' performance and in evaluating instructional programs, a variety of technical questions related to curriculum-based measurement need to be investigated. A first, critical question asks from what curriculum material it is appropriate to create the measurement task. Within any given curriculum sequence, a decision must be made regarding the level of difficulty from which the stimulus materials will be sampled. In reading, for example, words and passages may be selected from instructional level, independent level, or frustration level (Mirkin & Deno, 1979). Intuitively, it would seem that material used for continuous performance assessment should be neither too difficult (frustration level) nor too easy (independent level). If too difficult, performance will be low and the rate of increase too slow, thus precluding the use of data for evaluating the effects of instruction. Conversely, with the use of very easy material, very fast growth may occur in which student performance reaches a ceiling. In that event, changes must be made in the test stimulus for additional growth to be shown. If so, it will be difficult to determine the effect of a change in instructional strategy since it will be confounded with change in the test stimuli.

Somewhere between the two difficulty extremes, stimulus material must be identified that may be used over a relatively long time period to reliably and sensitively monitor student progress and reflect the effects of changes in the instructional program. When repeatedly measuring student performance over time in this manner, the measurement items will have to be kept constant, since changes in the testing procedures (e.g., items) would be confounded with changes in the instruc-

tional program (Campbell & Stanley, 1963). For this reason, direct measurement of student performance on the daily instructional task, as often recommended (Lovitt, 1967), may not be a workable solution to the question of what to measure in the curriculum.

The purpose of this study was to investigate how item selection in curriculum-based reading measurement impacts several technical characteristics of the measurement system. The measurement procedures investigated were based on research by Deno, Mirkin, and Chiang (1981) that established the validity of reading aloud from basal text vocabulary for measuring reading achievement. A first concern was how the population of words from which items were sampled for daily measurement influences the level and variability of student performance. A second question was how the size of the population of words influences the sensitivity of the daily measurement system for evaluating instructional programs. To investigate these issues, three measures were developed, differing only with respect to the size of the population of vocabulary words from which test items for daily testing were sampled.

Method

Subjects

Five special education resource teachers in the Minneapolis Public Schools, who had volunteered to participate in the study, were asked to list students who were reading at the second, third, or fourth grade instructional level. Four students were randomly selected from each teacher's list; these 20 students served as subjects in the study.

Materials

To develop daily measures of the student's reading performance, the

following procedures were used.

First, for each student three populations of reading vocabulary words were created using the Harris-Jacobson Word List (1972). The first and largest population, called Across-Grade list (AG), consisted of the entire pool of Harris-Jacobson words from the Preprimer-Grade 1 through Grade 4. A second population, termed the Grade-Level list (GL), consisted of all the Harris-Jacobson words from within the student's grade level. The third, Instructional-Level list (IL), was a subset of 200 words drawn at random from the GL population. The three populations differed, then, in terms of the scope of reading vocabulary words included. The scope of the AG population was the largest and the scope of the IL population was the smallest.

Daily word lists for testing were then created by drawing 60 words at random from each of the three populations. A different random sample from the respective domains was drawn each day to compose the daily test. Twenty word lists for each domain were created by random sampling with replacement. Therefore, the amount of repetition (words appearing more than once) from day to day within each list increased considerably from the Across-Grade list to the Instructional-Level list. Each teacher was given a set of 20 of each type of word list for every student.

Procedures

To determine an appropriate Grade-Level list in which to place the child, the student read aloud from the Grade-Level word lists for grades 1, 2, 3, and 4 for 30 seconds each. This procedure was repeated for five days. During this period, the teachers gave the Grade-Level

word list reading tests without specific instruction on the words. The number of words read correctly and incorrectly on each of the four Grade-Level word lists was recorded daily, and the student was placed for instruction in the Grade-level population where the median number of words read correctly over the five days was the highest.

Beginning the second week, the teachers initiated instruction for all their students. Each teacher was given the 200 Instructional-Level words that were drawn from the Grade-Level population in which the child had been placed. Each student was instructed individually for 10 minutes daily on the 200 word Instructional-Level set.

Immediately following the instructional period the student took a 30-second word reading test on each of the three populations of words using the daily test lists that had been created. The number of words read correctly and incorrectly on each type of word list was recorded by the teacher, and three daily performance graphs were created displaying correct and incorrect word reading.

Throughout the course of the study, the performance graphs were evaluated to determine the amount of improvement in the student's reading performance. Decisions were made weekly regarding whether to change a student's program. Attempts were made to incorporate procedures specific to that student's graphed performance (e.g., if a student's error rate was high, a change might be made to include error correction or a response cost procedure to reduce errors). In the event that five days of data were insufficient to reveal clear performance trends, the previous interventions were continued for two days and the judgment decision process resumed after the seventh day. A maximum of 15 days was allowed for keeping the same instructional format. When a decision

to change was made, the instructional intervention was implemented for another five days, after which the above procedure was applied again.

Results

Two primary analyses were conducted to assess the influence of the different word populations on the measurement data. The first analysis addressed the effects on the sensitivity of each test procedure to growth, and to variability in performance. The second analysis was conducted to assess the effects of the different item populations on evaluating changes in the instructional program.

Differences in Measurement Characteristics

Analysis of student performance on each type of daily test, from pretest to posttest, indicated differential sensitivity as a function of the population from which the daily test was created (see Table 1). When the populations were compared with respect to the mean difference in number of words read correct from pretest to posttest (i.e., the mean of the last three days), a reliable difference was found between populations, with the difference greatest for the Instructional-Level lists, followed by the Grade-Level lists and the least gain occurring with the Across-Grade lists. When accuracy was analyzed, a greater gain in percent of words read correct was obtained on the Grade-Level lists than on the Across-Grade lists. In this analysis, however, no reliable difference in gain was obtained between the Instructional-Level and Grade-Level lists.

Insert Table 1 about here

In Table 2, the semi-interquartile range (i.e., one-half the difference between the 75th and 25th percentile scores) is presented for the word list tests drawn from each population. These semi-interquartile ranges are presented for the fourth and twelfth instructional days. As can be seen, the obtained semi-interquartile ranges for all list scores remained quite consistent from Day 4 to Day 12. Differences in variability between lists also were quite consistent and small, with the GL lists the smallest on both Days 4 and 12. Variability also can be contrasted by examining the standard deviations presented in Table 1; the variability of the scores was consistently smaller on the tests created from the Across-Grade population of items.

Insert Table 2 about here

Differences in Evaluation of Instruction

A second analysis addressed the question of how well the measures created from each population could be used to evaluate changes in the instructional programs. To do so, the graphs of all students (3 per student) were randomly placed in folders (60 graphs). The graphs were presented independently to four judges; folders allowed the judges to see only the student's actual performance, with no information regarding type of word list, or even scaling of the axes.

Each judge was instructed to examine the student's performance in relation to the introduction of new instructional strategies, and decide whether the intervention had an effect upon the student's performance. Judges were told to attend both to the number of words read correct and incorrect. An effect on performance was defined when the number of

words read correct increased and errors stayed the same or correct stayed the same but errors decreased. Effects also were to include instances involving an increase in errors with no change in correct or a decrease in number of words read correct. Judges were instructed to attend to variability of performance, along with increases/decreases in either corrects or errors. One further definitional aspect involved the magnitude of change, that is, how much of an increase or decrease was needed to occur for an effect to be judged. An arbitrary value of 2 to 3 words was used as the magnitude sufficient to consider an effect.

When the above definitional standards for judging whether an intervention has an effect were used, a coefficient of concordance (agreements divided by agreements + disagreements) of .67 was attained for the first two judges; for judges three and four, a coefficient of .63 was obtained. While these coefficients are relative low, they are consistent with other published reports of this type of analysis.

In Table 3, the percents of interventions deemed to have an effect are presented for each word population and each judge. Table 4 presents the combined results for judges 1 and 2 and the combined results of judges 3 and 4. List population size was related reliably to the number of treatment changes judged effective by judges 1 and 2, both separately and combined. Chi square analysis revealed that judges 1 and 2 identified the lowest percent of apparent effects on AG scores ($X^2 = 6.2$ for judges 1 and 2; $X^2 = 6.02$ for judge 1; $X^2 = 5.96$ for judge 2). When the same analysis was conducted using judges 3 and 4, this finding was not replicated. For judges 3 and 4, no reliable effects for different list populations were obtained.

Insert Tables 3 and 4 about here

Discussion

The issue of item population is an important consideration in the development of a curriculum-based evaluation system. Not only might student performance vary as a function of the population from which samples are drawn, but the utility of using this type of data to evaluate instructional programs also may be influenced by the population size. It appears from this study that the best measurement system will be comprised of items which sample from the grade level at which the student is functioning although not necessarily the material in which instruction is being given. When a measure of student performance included words from the grade level in which the student was placed, growth was more rapid than when the words were drawn from a broader population of words. Further, weak evidence was obtained that the measurement systems based on the grade-level populations would produce performance graphs that would contribute more clearly to a visual analysis of instructional effects. The lack of consistency in judges' use of the graphs of student performance to evaluate instruction is perplexing, however. One can only assume that either the judges were inadequately trained, that the instructional interventions were not sufficiently powerful, or that the measures were not adequately sensitive. At present, no basis exists for selecting one explanation rather than another.

One important finding of the study was that a daily measurement system may be developed for reading instruction that can be used over an

extended period of time without having to be revised or changed. To create such a system requires specifying a broad enough population so that a ceiling is not obtained and narrow enough so it is sensitive to performance change. The effect of this implication is that measurement and instruction can proceed in a complementary fashion without undue domination by either. Teachers need not teach to the test nor limit their testing to instructional units.

References

Campbell, D. T., & Stanley, J. C. Experimental and quasi-experimental designs for research. Chicago: Rand McNally, 1963.

Carver, R. P. Two dimensions of tests: Psychometric and edumetric. American Psychologist, 1974, 29, 512-518.

Deno, S., Mirkin, P., & Chiang, B. Identifying valid measures of reading for use in continuous evaluation of IEP goals. Exceptional Children, in press.

Harris, A. J., & Jacobson, M. D. Basic elementary reading vocabularies - The first R series. New York: MacMillan, 1972.

Jenkins, J., & Pany, D. Standardized achievement tests: How useful for education? Exceptional Children, 1978, 44(6), 448-453.

Lovitt, T. Assessment of children with learning disabilities. Exceptional Children, 1967, 34(4), 233-239.

Lovitt, T., Schaff, M., & Sayre, E. The use of direct and continuous measurement to evaluate reading materials and pupil performance. Focus on Exceptional Children, 1970, 2, 1-11.

Mirkin, P. K., & Deno, S. L. Formative evaluation in the classroom: An approach to improving instruction (Research Report 10). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities, 1979.

Salvia, J., & Ysseldyke, J. E. Assessment in special and remedial education (2nd ed.). Boston: Houghton Mifflin, 1981.

Skager, R. The system for objectives-based evaluation. Reading, Evaluation Comment, 1971, 3(1), 6-11.

White, O., & Haring, N. Exceptional teaching (2nd ed.). New York: Charles Merrill, 1980.

Table 1

Mean Gain in Performance from Pre to Posttest[a]

Measure	List IL	GL	AG
Number of Words[b]	5.38 (3.50)	3.73 (3.59)	1.25 (2.98)
Percent of Words[c]	15.20 (13.53)	19.68 (17.47)	9.05 (12.19)

[a]Entries are the means and standard deviations (in parentheses) of the differences between pretest and posttest.

[b]Mean gain for the AG list was significantly different ($p < .05$) from that on either the IL or GL lists.

[c]Mean gain for the GL list was significantly different ($p < .05$) than on the AG list.

Table 2

Semi-Interquartile Ranges of Performance on Each Test List

Word List	Day 4	Day 12
Instructional-Level	4.10	4.05
Grade-Level	3.33	2.99
Across-Grade	4.68	5.18

Table 3

Percentages of Interventions Judged to have Apparent Effects, Not Apparent Effects, or Not Enough Data as a Function of Judges and List Population[a]

List Population	Effects Apparent				Effects Not Apparent				Not Enough Data			
	Jdg 1	Jdg 2	Jdg 3	Jdg 4	Jdg 1	Jdg 2	Jdg 3	Jdg 4	Jdg 1	Jdg 2	Jdg 3	Jdg 4
IL	42	51	65	49	20	16	24	29	38	33	11	22
GL	51	61	53	53	40	32	38	26	9	8	9	21
AG	29	36	62	51	37	29	28	28	34	35	10	21

[a]Difference between three list populations was significant ($p < .05$).

Table 4

Percentages of Interventions Judged to have Apparent Effects, Not Apparent Effects, and Not Enough Data by Judges 1 and 2 Combined and Judges 3 and 4 Combined[a]

List Population	Effects Apparent		Effects Not Apparent		Not Enough Data	
	Jdgs 1&2[b]	Jdgs 3&4	Jdgs 1&2	Jdgs 3&4	Jdgs 1&2	Jdgs 3&4
IL	46	58	18	27	36	15
GL	56	53	36	30	8	17
AG	32	57	33	28	35	15

[a]Interrater reliability was .67 for Judges 1 and 2 and .63 for Judges 3 and 4.

[b]Difference between three list populations was significant ($p < .05$).

PUBLICATIONS

Institute for Research on Learning Disabilities
University of Minnesota

The Institute is not funded for the distribution of its publications. Publications may be obtained for $3.00 per document, a fee designed to cover printing and postage costs. Only checks and money orders payable to the University of Minnesota can be accepted. All orders must be prepaid.

Requests should be directed to: Editor, IRLD, 350 Elliott Hall; 75 East River Road, University of Minnesota, Minneapolis, MN 55455.

Ysseldyke, J. E. Assessing the learning disabled youngster: The state of the art (Research Report No. 1). November, 1977.

Ysseldyke, J. E., & Regan, R. R. Nondiscriminatory assessment and decision making (Monograph No. 7). February, 1979.

Foster, G., Algozzine, B., & Ysseldyke, J. Susceptibility to stereotypic bias (Research Report No. 3). March, 1979.

Algozzine, B. An analysis of the disturbingness and acceptability of behaviors as a function of diagnostic label (Research Report No. 4). March, 1979.

Algozzine, B., & McGraw, K. Diagnostic testing in mathematics: An extension of the PIAT? (Research Report No. 5). March, 1979.

Deno, S. L. A direct observation approach to measuring classroom behavior: Procedures and application (Research Report No. 6). April, 1979.

Ysseldyke, J. E., & Mirkin, P. K. Proceedings of the Minnesota round-table conference on assessment of learning disabled children (Monograph No. 8). April, 1979.

Somwaru, J. P. A new approach to the assessment of learning disabilities (Monograph No. 9). April, 1979.

Algozzine, B., Forgnone, C., Mercer, C. D., & Trifiletti, J. J. Toward defining discrepancies for specific learning disabilities: An analysis and alternatives (Research Report No. 7). June, 1979.

Algozzine, B. The disturbing child: A validation report (Research Report No. 8). June, 1979.

Note: Monographs No. 1 - 6 and Research Report No. 2 are not available for distribution. These documents were part of the Institute's 1979-1980 continuation proposal, and/or are out of print.

Ysseldyke, J. E., Algozzine, B., Regan, R., & Potter, M. Technical adequacy of tests used by professionals in simulated decision making (Research Report No. 9). July, 1979.

Jenkins, J. R., Deno, S. L., & Mirkin, P. K. Measuring pupil progress toward the least restrictive environment (Monograph No. 10). August, 1979.

Mirkin, P. K., & Deno, S. L. Formative evaluation in the classroom: An approach to improving instruction (Research Report No. 10). August, 1979.

Thurlow, M. L., & Ysseldyke, J. E. Current assessment and decision-making practices in model programs for the learning disabled (Research Report No. 11). August, 1979.

Deno, S. L., Chiang, B., Tindal, G., & Blackburn, M. Experimental analysis of program components: An approach to research in CSDC's (Research Report No. 12). August, 1979.

Ysseldyke, J. E., Algozzine, B., Shinn, M., & McGue, M. Similarities and differences between underachievers and students labeled learning disabled: Identical twins with different mothers (Research Report No. 13). September, 1979.

Ysseldyke, J., & Algozzine, R. Perspectives on assessment of learning disabled students (Monograph No. 11). October, 1979.

Poland, S. F., Ysseldyke, J. E., Thurlow, M. L., & Mirkin, P. K. Current assessment and decision-making practices in school settings as reported by directors of special education (Research Report No. 14). November, 1979.

McGue, M., Shinn, M., & Ysseldyke, J. Validity of the Woodcock-Johnson psycho-educational battery with learning disabled students (Research Report No. 15). November, 1979.

Deno, S., Mirkin, P., & Shinn, M. Behavioral perspectives on the assessment of learning disabled children (Monograph No. 12). November, 1979.

Sutherland, J. H., Algozzine, B., Ysseldyke, J. E., & Young, S. What can I say after I say LD? (Research Report No. 16). December, 1979.

Deno, S. L., & Mirkin, P. K. Data-based IEP development: An approach to substantive compliance (Monograph No. 13). December, 1979.

Ysseldyke, J., Algozzine, B., Regan, R., & McGue, M. The influence of test scores and naturally-occurring pupil characteristics on psycho-educational decision making with children (Research Report No. 17). December, 1979.

Algozzine, B., & Ysseldyke, J. E. Decision makers' prediction of students' academic difficulties as a function of referral information (Research Report No. 18). December, 1979.

Ysseldyke, J. E., & Algozzine, B. Diagnostic classification decisions as a function of referral information (Research Report No. 19). January, 1980.

Deno, S. L., Mirkin, P. K., Chiang, B., & Lowry, L. Relationships among simple measures of reading and performance on standardized achievement tests (Research Report No. 20). January, 1980.

Deno, S. L., Mirkin, P. K., Lowry, L., & Kuehnle, K. Relationships among simple measures of spelling and performance on standardized achievement tests (Research Report No. 21). January, 1980.

Deno, S. L., Mirkin, P. K., & Marston, D. Relationships among simple measures of written expression and performance on standardized achievement tests (Research Report No. 22). January, 1980.

Mirkin, P. K., Deno, S. L., Tindal, G., & Kuehnle, K. Formative evaluation: Continued development of data utilization systems (Research Report No. 23). January, 1980.

Deno, S. L., Mirkin, P. K., Robinson, S., & Evans, P. Relationships among classroom observations of social adjustment and sociometric rating scales (Research Report No. 24). January, 1980.

Thurlow, M. L., & Ysseldyke, J. E. Factors influential on the psychoeducational decisions reached by teams of educators (Research Report No. 25). February, 1980.

Ysseldyke, J. E., & Algozzine, B. Diagnostic decision making in individuals susceptible to biasing information presented in the referral case folder (Research Report No. 26). March, 1980.

Thurlow, M. L., & Greener, J. W. Preliminary evidence on information considered useful in instructional planning (Research Report No. 27). March, 1980.

Ysseldyke, J. E., Regan, R. R., & Schwartz, S. Z. The use of technically adequate tests in psychoeducational decision making (Research Report No. 28). April, 1980.

Richey, L., Potter, M., & Ysseldyke, J. Teachers' expectations for the siblings of learning disabled and non-learning disabled students: A pilot study (Research Report No. 29). May, 1980.

Thurlow, M. L., & Ysseldyke, J. E. Instructional planning: Information collected by school psychologists vs. information considered useful by teachers (Research Report No. 30). June, 1980.

Algozzine, B., Webber, J., Campbell, M., Moore, S., & Gilliam, J. Classroom decision making as a function of diagnostic labels and perceived competence (Research Report No. 31). June, 1980.

Ysseldyke, J. E., Algozzine, B., Regan, R. R., Potter, M., Richey, L., & Thurlow, M. L. Psychoeducational assessment and decision making: A computer-simulated investigation (Research Report No. 32). July, 1980.

Ysseldyke, J. E., Algozzine, B., Regan, R. R., Potter, M., & Richey, L. Psychoeducational assessment and decision making: Individual case studies (Research Report No. 33). July, 1980.

Ysseldyke, J. E., Algozzine, B., Regan, R., Potter, M., & Richey, L. Technical supplement for computer-simulated investigations of the psychoeducational assessment and decision-making process (Research Report No. 34). July, 1980.

Algozzine, B., Stevens, L., Costello, C., Beattie, J., & Schmid, R. Classroom perspectives of LD and other special education teachers (Research Report No. 35). July, 1980.

Algozzine, B., Siders, J., Siders, J., & Beattie, J. Using assessment information to plan reading instructional programs: Error analysis and word attack skills (Monograph No. 14). July, 1980.

Ysseldyke, J., Shinn, M., & Epps, S. A comparison of the WISC-R and the Woodcock-Johnson Tests of Cognitive Ability (Research Report No. 36). July, 1980.

Algozzine, B., & Ysseldyke, J. E. An analysis of difference score reliabilities on three measures with a sample of low achieving youngsters (Research Report No. 37). August, 1980.

Shinn, M., Algozzine, B., Marston, D., & Ysseldyke, J. A theoretical analysis of the performance of learning disabled students on the Woodcock-Johnson Psycho-Educational Battery (Research Report No. 38). August, 1980.

Richey, L. S., Ysseldyke, J., Potter, M., Regan, R. R., & Greener, J. Teachers' attitudes and expectations for siblings of learning disabled children (Research Report No. 39). August, 1980.

Ysseldyke, J. E., Algozzine, B., & Thurlow, M. L. (Eds.). A naturalistic investigation of special education team meetings (Research Report No. 40). August, 1980.

Meyers, B., Meyers, J., & Deno, S. Formative evaluation and teacher decision making: A follow-up investigation (Research Report No. 41). September, 1980.

Fuchs, D., Garwick, D. R., Featherstone, N., & Fuchs, L. S. On the determinants and prediction of handicapped children's differential test performance with familiar and unfamiliar examiners (Research Report No. 42). September, 1980.

Algozzine, B., & Stoller, L. Effects of labels and competence on teachers' attributions for a student (Research Report No. 43). September, 1980.

Ysseldyke, J. E., & Thurlow, M. L. (Eds.). The special education assessment and decision-making process: Seven case studies (Research Report No. 44). September, 1980.

Ysseldyke, J. E., Algozzine, B., Potter, M., & Regan, A. A descriptive study of students enrolled in a program for the severely learning disabled (Research Report No. 45). September, 1980.

Marston, D. Analysis of subtest scatter on the tests of cognitive ability from the Woodcock-Johnson Psycho-Educational Battery (Research Report No. 46). October, 1980.

Algozzine, B., Ysseldyke, J. E., & Shinn, M. Identifying children with learning disabilities: When is a discrepancy severe? (Research Report No. 47). November, 1980.

Fuchs, L., Tindal, J., & Deno, S. Effects of varying item domain and sample duration on technical characteristics of daily measures in reading (Research Report No. 48). January, 1981.

Marston, D., Lowry, L., Deno, S. & Mirkin, P. An analysis of learning trends in simple measures of reading, spelling, and written expression: A longitudinal study (Research Report No. 49). January, 1981.

Marston, D., & Deno, S. The reliability of simple, direct measures of written expression (Research Report No. 50). January, 1981.

Epps, S., McGue, M., & Ysseldyke, J. E. Inter-judge agreement in classifying students as learning disabled (Research Report No. 51). February, 1981.

Epps, S., Ysseldyke, J. E., & McGue, M. Differentiating LD and non-LD students: "I know one when I see one" (Research Report No. 52). March, 1981.

Evans, P. R., & Peham, M. A. S. Testing and measurement in occupational therapy: A review of current practice with special emphasis on the Southern California Sensory Integration Tests (Monograph No. 15). April, 1981.

Fuchs, L., Wesson, C., Tindal, G., & Mirkin, P. Teacher efficiency in continuous evaluation of IEP goals (Research Report No. 53). June, 1981.

Fuchs, D., Featherstone, N., Garwick, D. R., & Fuchs, L. S. The importance of situational factors and task demands to handicapped children's test performance (Research Report No. 54). June, 1981.

Tindal, G., & Deno, S. L. Daily measurement of reading: Effects of varying the size of the item pool (Research Report No. 55). July 1981.